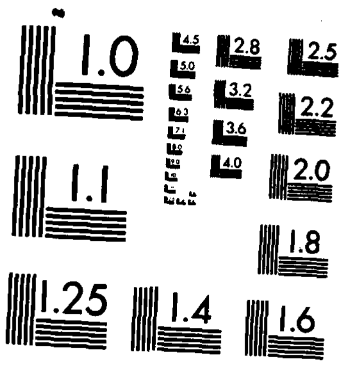
AD-A164 087 THE METHOD OF SIEVES A SURVEY OF RECENT APPLICATIONS 1/1
(U) FLORIDA STATE UNIV AALLAHASSEE DEPT OF STATISTICS
I W MCKEAGUE DEC 85 FSU-STATISTICS-M718 ARO-19367.34-MA

UNCLASSIFIED DAAG29-82-K-0168 F/G 12/1 NL



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER ARO 19367.34-MA	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) The Method of Sieves: A Survey of Recent Applications		5. TYPE OF REPORT & PERIOD COVERED Technical Report
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Ian W. McKeague		8. CONTRACT OR GRANT NUMBER(s) DAAG29-82-K-0168
9. PERFORMING ORGANIZATION NAME AND ADDRESS Florida State University		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS U. S. Army Research Office Post Office Box 12211 Research Triangle Park, NC 27709		12. REPORT DATE December, 1985
		13. NUMBER OF PAGES 9
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) Unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

DTIC ELECTE FEB 1 1 1986 S D B

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

The view, opinions, and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy, or decision, unless so designated by other documentation

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Method of Sieves
Nonparametric Estimation
Estimators
Subsets
Sample Size
Least Squares Estimators

DTIC FILE COPY

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

The method of sieves is a technique of nonparametric estimation in which estimators are restricted by an increasing sequence of subsets of the parameter space with the subsets indexed by the sample size. The need for this

DD FORM 1473, 1 JAN 73 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

(2)

20. ABSTRACT CONTINUED

technique arises in situations where the parameter space is too large for the existence or consistency of unconstrained maximum likelihood or least squares estimators. Grenander [10] developed the abstract theory of the method of sieves and provided a wealth of examples illustrating its use.

ARO 19367.34-MA

THE METHOD OF SIEVES: A SURVEY OF
RECENT APPLICATIONS[1]

by

Ian W. McKeague[2]

Florida State University
Department of Statistics
Tallahassee, Florida 32306

FSU Statistics Report M-718
USARO Technical Report No. D-84

December, 1985

[1] This article is to appear under the entry SIEVES, METHOD OF in the Encylopedia of Statistical Sciences, published by John Wiley and Sons.

[2] Research supported by the U.S. Army Research Office under Grant DAAG 29-82-K-0168.

INTRODUCTION

The method of sieves is a technique of nonparametric estimation in which estimators are restricted by an increasing sequence of subsets of the parameter space with the subsets indexed by the sample size. The need for this technique arises in situations where the parameter space is too large for the existence or consistency of unconstrained maximum likelihood or least squares estimators. Grenander [10] developed the abstract theory of the method of sieves and provided a wealth of examples illustrating its use.

Geman and Hwang [9] have shown that the method of sieves leads to consistent nonparametric estimators in very general settings. In practice, the sequence of subsets of the parameter space (which comprise the sieve) needs to be carefully chosen to exploit the specific structural properties of the problem. It would be desirable to know how to construct the sieve to yield an optimal rate of convergence of the sieve estimator, but this question has only been studied in some special cases.

In some nonparametric problems, typically where a monotonicity condition holds, the method of maximum likelihood is directly applicable without the need for a sieve. For instance, under monotonicity of the probability density function the maximum likelihood estimator, based on an iid sample, exists and is consistent in L^1-norm, see Grenander

[10, p. 402]. Similar results have been obtained for monotone failure rate functions [15], and unimodal densities [19]. However without order restrictions the direct method of maximum likelihood usually fails in nonparametric problems. The method of sieves then presents itself as one of several alternative approaches, others being the method of penalized maximum likelihood*, orthogonal series methods, kernel methods*, spline methods and the Bayesian approach. These techniques are themselves closely related to the method of sieves, see the discussion on this matter in Grenander [10, p. 7] and Geman and Hwang [9, p. 403]. The distinguishing feature of the method of sieves is that it makes use of an optimization principle subject to constraints which depend on the sample size. The following examples of the method of sieves supplement those already mentioned under the entry METHOD OF SIEVES.

TRANSLATE OF WIENER PROCESS

Let $W(t)$, $t \geq 0$ be a standard Wiener process* and α an unknown function of $t \in [0,1]$. Suppose that n independent identically distributed (iid) copies X_i, $i = 1, \ldots, n$ of the signal + noise process

$$X(t) = \int_0^t \alpha(s)ds + W(t), \; t \in [0,1] \qquad (1)$$

are observed. The parameter space for this problem is

$L^2[0,1]$, the space of square integrable functions on $[0,1]$. Grenander [10, p. 424] considered a sieve of the form

$$S_{d_n} = \left\{\alpha(t):\ \alpha(t) = \sum_{r=1}^{d_n} \alpha_r \phi_r(t)\right\} \tag{2}$$

where $(\phi_r,\ r \ge 1)$ is a complete orthonormal sequence in $L^2[0,1]$. The maximum likelihood estimator contained in S_{d_n} is given by

$$\hat{\alpha}^{(n)}(t) = \sum_{r=1}^{d_n} \hat{\alpha}_r^{(n)} \phi_r(t) \tag{3}$$

where

$$\hat{\alpha}_r^{(n)} = \frac{1}{n} \sum_{i=1}^{n} \int_0^1 \phi_r(t)\,dX_i(t).$$

It can be shown that $\hat{\alpha}^{(n)}$ is consistent in L^2-norm as $n \to \infty$ provided $d_n \uparrow \infty$ and $d_n/n \to 0$, see Nguyen and Pham [18] and McKeague [17]. The estimator (3) was first studied by Ibragimov and Khasminskii [11] who defined it from a point of view suggested by Cencov's [5] method of orthogonal series for density estimation*. Ibragimov and Khasminskii showed that within the parameter space of Lipschitz functions of order γ, $0 < \gamma \le 1$, the estimator $\hat{\alpha}^{(n)}$ can be designed to attain the optimal rate of convergence (in the sense of an asymptotic minimax property) over all estimators. The optimal rate of convergence of the mean square error is $O(n^{-2\gamma/(2\gamma+1)})$ and this can be achieved by using the Fourier sieve

$$S_{d_n} = \left\{\alpha: \alpha(t) = \sum_{r=-d_n}^{d_n} \alpha_r e^{2\pi irt}\right\}$$

with $d_n = [n^{1/(2\gamma+1)}]$, where [ ] denotes the integer part.

Another sieve for this problem is given by

$$S_{m_n} = \left\{\alpha \in L^2[0,1]: \sum_{r=1}^{\infty} a_r^2 <\alpha, \phi_r>^2 \le m_n\right\},$$

where < , > denotes the inner product in $L^2[0,1]$ and $\sum_{r\ge 1} a_r^{-2} < \infty$. This sieve has been studied by Geman and Hwang [9], and Antoniadis [3] for general Gaussian processes. Antoniadis showed that this sieve estimator is consistent provided $m_n \uparrow \infty$ and $m_n = O(n^{1-\varepsilon})$ for some $\varepsilon > 0$. Beder [22] has studied sieves of the form (2) for general Gaussian processes. Other approaches to the problem can be found in [14, 16, 20].

INTENSITY OF A POINT PROCESS

Let $N(t)$, $t \ge 0$ be a point process* with intensity

$$\lambda(t) = \sum_{j=1}^{p} \alpha_j(t) Y_j(t) \tag{4}$$

where $\alpha_1, \ldots, \alpha_p$ are unknown functions and $Y_1, \ldots, Y_p$ are observable covariate processes. Practical examples of this model arise in reliability and biomedical settings. For instance, suppose that a subject has been exposed to p carcinogens. Let X be the time of the initial detection of cancer. Then a plausible "competing risks model" for

the hazard function $\lambda(t)$ of X is given by (4) where $\alpha_1, \ldots, \alpha_p$ represent the changes in the relative hazard rates of the p carcinogens with age and $Y_j(t)$ is the cumulative exposure to the j^{th} carcinogen by age t. The model (4) was introduced by Aalen [1, 2] as an alternative to the proportional-hazard regression model* of Cox [7]. Aalen introduced an estimator of the integrated hazard functions $\int_0^t \alpha_j(s)ds$.

The method of sieves is able to provide estimators of the α_j's themselves. Suppose that n iid copies of the processes $N(t)$, $Y_j(t)$ are observed over $[0,1]$. In the case $p = 1$, Karr [1] used the sieve

$$S_{a_n} = \{\alpha \in L^1[0,1]: \alpha \text{ is absolutely continuous, } a_n \le \alpha \le a_n^{-1} \text{ and } |\alpha'| \le a_n^{-1}\alpha\},$$

and showed that the maximum likelihood estimator of α_1 restricted by this sieve is strongly consistent in L^1-norm where $a_n = n^{-\frac{1}{4}+\eta}$, with $0 < \eta < \frac{1}{4}$. For models with more than one covariate McKeague [17] has used the orthogonal series sieve (2) to obtain consistent estimators of $\alpha_1, \ldots, \alpha_p$ for a general semimartingale regression model which contains the point process model (4) and diffusion process models as special cases.

STATIONARY PROCESSES

Some recent applications of the method of sieves have been motivated by problems in the area of engineering known as system identification. A stationary process is observed over a long period of time and the engineer seeks to reconstruct the "black box" which produced the process. In practice this amounts to estimation of a spectral density or a transfer function and similar considerations which led to the use of the method of sieves for probability density estimation are involved here.

Chow and Grenander [6] consider estimation of the spectral density of a stationary Gaussian process $\{X_t, t = 1,2,\ldots\}$ with mean zero and covariance $r_t = E(X_s X_{s+t}) = \int_{-\pi}^{\pi} e^{it\lambda} f(\lambda)d\lambda$, where f is the spectral density. They employ a sieve of the form

$$S_{\mu_n} = \left\{f\colon\ f = 1/g \text{ and } \int_{-\pi}^{\pi} \left(\frac{d}{d\lambda}\, g\right)^2 d\lambda \le \frac{1}{\mu_n}\right\},$$

where n is the length of observation time. They show that an approximate maximum likelihood estimator of f restricted to S_{μ_n} is strongly consistent in $L^1[-\pi,\pi]$ provided $\mu_n = n^{-(1-\delta)}$ where $0 < \delta < 1$.

Ljung and Yuan [13] consider the problem of estimating the transfer function of a linear stochastic system given by

$$y(t) = \sum_{k=1}^{\infty} g_k u(t-k) + w(t), \quad t = 1,2,\ldots .$$

Here u(t) and y(t) are the input and output, respectively, at time t and {w(t)} is supposed to be a stationary process. A reasonable sieve for the transfer function $h(\omega) = \sum_{k=1}^{\infty} g_k e^{-ik\omega}$, $\omega \in [-\pi,\pi]$, is given by

$$S_{d_n} = \left\{h:\ h(\omega) = \sum_{k=1}^{d_n} g_k e^{-ik\omega}\right\},$$

where n is the length of observation time of input and output processes. The results of Ljung and Yuan show that the sieve estimator, formed by using the least squares estimates of $g_1, \ldots, g_{d_k}$, is uniformly consistent provided $d_n = [n^\alpha]$, $0 < \alpha < \frac{1}{4}$.

Bagchi [4] has used the method of sieves to estimate the distributed delay function α of the following linear time-delayed system:

$$dX_t = \left[\int_{-b}^{0} \alpha(u) X_{t+u} du\right] dt + dW_t,$$

where $\{W_t, -\infty < t < \infty\}$ is a standard Wiener process. The sieve is given by

$$S_{d_T} = \left\{\alpha \in L^2[-b,0]:\ \alpha(u) = \sum_{r=1}^{d_T} \alpha_r \phi_r(u)\right\},$$

where $(\phi_r, r \geq 1)$ is a complete orthonormal sequence in $L^2[-b,0]$ and T is the length of observation time of the process X. Bagchi shows that the maximum likelihood

estimator restricted to S_{d_T} is consistent in L^2-norm provided $d_T \uparrow \infty$ and $d_T^2/T \to 0$ as $T \to \infty$.

REFERENCES

[1] Aalen, O.O. (1978). Ann. Statist., 6, 701-726.

[2] Aalen, O.O. (1980). Lect. Notes Statist., 2, 1-25. Springer-Verlag, New York.

[3] Antoniadis, A. (1985). Institut de Mathématiques Appliquées de Grenoble Technical Report.

[4] Bagchi, A. (1985). Systems and Control Letters, 5, 339-345.

[5] Cencov, N.N. (1962). Soviet Math., 3, 1559-1562.

[6] Chow, Y.-S. and Grenander, U. (1985). Ann. Statist., 13, 998-1010.

[7] Cox, D.R. (1972). J. R. Statist. Soc. B., 34, 187-220.

[8] Geman, S. (1980). Colloquia Mathematica Societatis János Bolyai, 32. Nonparametric Statistical Inference, Budapest (Hungary) 1980.

[9] Geman, S. and Hwang, C.-R. (1982). Ann. Statist., 10, 401-414.

[10] Grenander, U. (1981). Abstract Inference. Wiley, New York.

[11] Ibragimov, I.A. and Khasminskii, R.Z. (1981). Theory Prob. Appl., 25, 703-719.

[12] Karr, A.F. (1983). Center for Stochastic Processes,

Department of Statistics, U.N.C. Chapel Hill, Technical Report #46.

[13] Ljung, L. and Yuan, Z.-D. (1985). IEEE Trans. Automat. Contr., AC-30, 514-530.

[14] Mandelbaum, A. (1984). Ann. Statist., 12, 1448-1466.

[15] Marshall, A.W. and Proschan, F. (1965). Ann. Math. Statist., 36, 69-77.

[16] McKeague, I.W. (1984). Ann. Statist., 12, 1310-1323.

[17] McKeague, I.W. (1986). To appear in Ann. Statist., March 1986.

[18] Nguyen, H.T. and Pham, T.D. (1982). SIAM J. on Control and Optimization, 20, 603-611.

[19] Prakasa Rao, B.L.S. (1969). Sankhyā A, 31, 23-36.

[20] Tsirelson, B.S. (1982). Theory Prob. Appl., 27, 411-418.

[21] Walter, G.G. and Blum, J.R. (1984). Ann. Statist., 12, 372-379.

[22] Beder, J. (1985). A sieve estimator for the mean of a Gaussian process. Preprint.

END

FILMED

3-86

DTIC